**Article:**

Larsen, Rasmus Tolstrup, Christensen, Jan, Tang, Lars Hermann et al. (5 more authors) (2016) A SYSTEMATIC REVIEW AND META-ANALYSIS COMPARING CARDIOPULMONARY EXERCISE TEST VALUES OBTAINED FROM THE ARM CYCLE AND THE LEG CYCLE RESPECTIVELY IN HEALTHY ADULTS. International journal of sports physical therapy. pp. 1006-1039. ISSN 2159-2896

# A systematic review and meta-analysis comparing cardiopulmonary exercise test values obtained from the arm cycle and the leg cycle respectively in healthy adults

---

**Key words:** aerobic capacity, exercise testing, oxygen uptake, leg cycle, arm cycle, ergometer, systematic review, meta-analysis.

*Word Count: 2.738*

**Abstract (200 words)**

**Introduction**: The cardiopulmonary exercise test (CPET) assesses maximal oxygen uptake (VO2max) and is commonly performed on a leg cycle ergometer (LC). However, some individuals will rather perform the CPET on an arm cycle ergometer (AC).

**Objective**: To compare VO2max values obtained by the AC test and the LC test in healthy adults.

**Methods**: MEDLINE, EMBASE, CINAHL, and PEDro were searched in April 2015. Studies were included if they reported within comparison VO2max values obtained from CPET using AC and LC in healthy adults. The differences in VO2max ($ACLC_{diff}$) were pooled across studies using random effects meta-analysis and three different methods were used to estimate the ratio between the values obtained from the tests ($ACLC_{ratio}$).

**Results**: We included 41 studies with a total of 581 participants. The mean $ACLC_{diff}$ across studies was 12.5 ml/kg/min and 0.89 l/min with a mean $ACLC_{ratio}$ of 0.70. The $ACLC_{diff}$ was lower in studies with higher mean age and lower aerobic capacity.

**Conclusion**: There is linear association between the AC and LC values in healthy non-athletic individuals. The AC obtained values were on average 70% of the LC values. The magnitude of this difference appeared to be reduced in studies on older and less active populations.

## 1. Introduction

The cardiopulmonary exercise test (CPET) is the gold standard for the direct assessment of maximal oxygen uptake ($VO_{2max}$) [1-5]. $VO_{2max}$ determines the maximal ability for the human body to deliver, obtain and consume oxygen during maximal exercise and is a measure of maximum aerobic capacity [4]. Assessments of aerobic capacity are used by physicians and healthcare professionals to evaluate exercise capacity [5], exercise intolerance [6] and functional aerobic impairment [7], which all provide important information on health status and prognosis in various populations [2,8-11].

CPET is commonly performed on a treadmill or on a leg cycle ergometer (LC) [3,5]. However, due to disability, co-morbidity, preference or athletic discipline there is a need to investigate alternatives to LC [12]. In some cases, it could be more important to assess arm fitness when leg exercise is not feasible or possible [13-15]. A potential alternative is to perform the test with the upper body using an arm cycle ergometer (AC)[13]. The AC test is however challenged as studies have shown that untrained individuals will achieve a lower level of $VO_{2max}$ on the AC, due to a reduced stress on the cardiovascular system, compared to LC [12,15,16]. Having a smaller amount of muscle mass being active during the test, AC is likely to result in an earlier termination of the CPET due to peripheral factors such as an earlier onset of lactate threshold, rather than central cardiovascular limitations [12,17]. Whilst individual studies have directly assessed the difference in $VO2_{max}$ of a CPET conducted using AC compared to LC in healthy adults, we know of no previous systematic review of these studies.

The objectives of this study were to undertake a systematic review and meta-analysis of the $VO_{2max}$ achieved by AC compared to LC in healthy adults and to explore factors that may be predictive of this difference. The determination of this factor would allow the direct comparison of data obtained on the two tests.

## 2. Methods

This review was conducted and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [18].

### 2.1 Data sources and searches

Preliminary searches were conducted and relevant search terms identified. A formal search of the databases MEDLINE, EMBASE, CINAHL, and PEDro was undertaken in April 2015. References of the identified studies in the preliminary searches were screened and relevant search terms were added to the search strategy. The search strategy consisted of a combination of relevant keywords and MeSH/Thesaurus terms for: 1) direct assessment of $VO2_{max,}$ 2) a CPET performed on an AC and 3) a CPET performed on an LC. No language or publication limits were applied. The reference lists of identified studies were checked and we contacted the authors of unobtainable studies and evaluated papers suggested by experts in the field. Search strategies specified for MEDLINE is presented in appendix.

### 2.3 Study selection

Study selection was undertaken based on a priori defined criteria. Only original research papers reporting within comparison maximum or peak $VO_2$, as litres per minute (l/min) or as millilitre oxygen per kilogram per minute (ml/min/kg), were considered eligible for inclusion in this systematic review. The CPET had to be non-assisted on AC and LC.  We included studies in groups of healthy adults (age >18 years) with a reported level of physical activity < 300 minutes per week. People with higher physical activity levels were considered athletes and where therefore excluded [19].

Two authors (RTL, CK) independently screened titles and abstracts and assessed eligible articles in full-text. Any inconsistencies between authors were discussed and disagreement was solved by consultation of a third author (JC).

### 2.4 Data extraction and risk of bias assessment

The following information was extracted: sample size, gender distribution, mean age, mean height, body mass index (BMI) together with the $VO2_{max}$ values, peak respiratory exchange ratio (RER), CPET starting Watt, and Watt increment for both the AC and LC test.

*The Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies* [20] was used to assess the methodological quality of all included studies. Six items (6-10 and 13) were considered not applicable for the studies included in this review and thus did not contribute to the quality rating total score (SumQAT). Two authors (RTL and CK) independently extracted data and undertook the quality assessment. Inconsistencies between reviewers were discussed and in cases of disagreement, a third reviewer (JC) was consulted.

### 2.5 Data analysis

The mean $VO2_{max}$ difference between AC and LC ($ACLC_{diff}$) was calculated for each study. Given the within subject nature of these comparisons we adjusted the standard deviation of this difference for the within subject correlation using the method described in chapter 16.4.6.1 of the Cochrane Handbook [21]. The level of statistical heterogeneity was assessed using the $I^2$ score. The $ACLC_{diff}$, for ml/kg/min and l/min, were pooled across studies using a conservative random effects meta-analysis given the variation participant characteristics across included studies. Summary of the characteristics of included studies are expressed as median values and interquartile range (IQR).

We used meta-regressions to perform sub-group analyses to clarify, which variables were affecting the main analysis on the $ACLC_{diff.}$ The sub-groups included were: aerobic capacity (as a categorical variable based the Aastrand classification -“low”, “fair”, “average”, “good” or “high") [22], participant mean age (in years), participant gender (percentage of males), study risk of bias (SumQAT), and the difference in peak RER values during test.

Three different approaches were used to find the ratio between AC and LC ($ACLC_{ratio}$). First a meta-analysis of the $ACLC_{ratio}$ was undertaken using the studies presenting the group mean ± standard deviation of the within comparison ratio (%). Second a linear regression model was determined using the group mean values. The linear regression analysis was weighted by sample size. Third the reported AC values were divided with the reported LC values, giving an estimate of the ratio in each study, which are expressed as a total mean ratio.

All analyses were performed using Review Manager 5.3 (Cochrane collaboration) software and Stata 14.0 software (StataCorp. 2013. *Stata Statistical Software: Release 14.9* College Station, TX: StataCorp LP). A p-value ≤ 0.05 was considered statistically significant.

## 3. Results

### 3.1 Study selection

Our database searches identified 3,300 records. After removing 617 duplicates, 2,683 unique studies remained. We excluded 2,510 studies by screening their title and abstract and 173 studies were considered eligible for full text review. Of these, 131 did not meet the inclusion criteria. Thus, 41 studies (published between 1973 and 2014) were included in the review [12,15,17,23-60]. Citations and reasons for full text exclusion are listed in appendix. The study selection process is summarised as a flow chart in Figure 1.

### 3.2 Description of studies

A summary of the characteristics of the included studies is provided in Table 1. The full characteristics of included studies are listed in appendix.

### 3.3 Risk of bias in included studies

Figure 2 presents a summary of the risk of bias in the included studies. The median SumQAT was 4 points, (IQR: 3 to 5). A detailed risk of bias of each study is listed in appendix.

#### *3.3.1 Research question and study population*

Although all included studies were judged to have a well-defined research question (item 1), 13 groups [15,17,25,30,34-38,44,47,52,53] had insufficient description of the study population (item 2). One study [43] described the participation rate of eligible subjects (item 3) and 13 [15,24,26,29,32,33,35,36,41,43,46,55,57] studies had a subject-recruitment within the same population (item 4). Four studies [15,38,39,46] included sample size justification (item 5).

*3.3.2 Outcome measures*

Five studies [17,34,40,45,53] did not report the $VO2_{max}$ as ml/kg/min but as l/min (item 11) and therefore not adjusting their outcome for subject weight.

*3.3.3 Blinding and statistical analysis*

One study [12] blinded the outcome assessor (item 12) and 12 studies [25,34,37,38,40,42,43,45,48,51,52,55] did not provide report a description of their statistical analysis methods (item 14).

**3.4 Meta-analysis of $VO2_{max}$difference between AC and LC**

A total of 36 groups (413 participants) reported data on the $ACLC_{diff}$ measured in ml/kg/min. The meta-analysis for the $ACLC_{diff}$ is shown in Figure 3. The pooled mean $VO_{2max}$ was 12.5 ml/kg/min, (95% CI: 10.3 to 14.7, $I^2$ = 59.9%, p > 0.001) higher for LC than AC. A total of 37 comparisons (415 participants) presented data of the $ACLC_{diff}$ in l/min with pooled mean $VO_{2max}$ of 0.89 l/min, (95% CI: 0.78 to 1.00, $I^2$ = 30.5%, p=0.043) higher for LC than AC as shown in figure 4.

**3.5 Subgroup analyses**

In univariable meta-regression and multivariable meta-regression, lower participant mean age and higher aerobic capacity were found to be significantly associated an increased $ACLC_{diff}$. The meta-regressions are shown in Table 2.

**3.6 Analyses of the AC/LC ratio**

The mean ratio between the AC and LC for the 37 groups (n=413 participants) reporting $VO2_{max}$ in ml/kg/min was 0.70 (95% CI: 0.66 to 0.73) in favour of the LC. The corresponding value of the 37 groups (n=415 participants) reporting $VO_{2max}$ in l/min, the mean $ACLC_{ratio}$ was 0.71 (95% CI: 0.66 to 0.75). The meta-analysis (n=46 studies) for the $ACLC_{ratio}$ across studies as 71%, (95% CI: 68 to 74, $I^2$ = 0%, p=0.530) (Figure 5). The coefficient for the linear regression between AC and LC mean VO2max was 0.65 ml/kg/min (95% CI: 0.48 to 0.81) with an $r^2$ of 0.689 (Figure 6).

## 4. Discussion

This systematic review and meta-analysis brings together data from 41 studies in 581 healthy individuals directly comparing $VO_{2max}$ values obtained from the AC compared to LC. We found the LC to have substantively higher $VO_{2max}$ value (mean difference: 12.5 ml/kg/min and 0.89 l/min) than AC. But with an $I^2$ value of 59.9% for the $ACLC_{diff}$ in ml/kg/min these results could be affected by substantial heterogeneity. Our results support the belief that the AC test achieves lower oxygen uptake values as it involves a smaller amount of muscle mass and places less stress on the cardiovascular system [12,15,16].

Both age and the aerobic capacity appear to be associated with the $ACLC_{diff.}$ The difference is decreased with increasing age and increased with better aerobic capacity. This was somehow expected, due to the fact that aerobic capacity decreases with age [22].

The RER represent the relationship between the volume of carbon dioxide and the volume of oxygen in every breath and it is recommended to continue $VO_{2max}$ tests until RER values above 1.1 are reached in order to obtain a valid CPET [23]. The majority of studies reporting RER values reported values in both tests to be above 1.1 [23,24,26-29,32,36,38,46,60]. Only one study reported RER values for the AC to be above 1.1 and RER values for the LC to be below 1.1 [23], and three studies reported RER for both test to be below 1.1 [33,39,49]. We expected the difference in the obtained RER values to affect the $ACLC_{diff}$. However, we did not find this relationship, which could be due to by a lack of power, as only 24 and 16 studies are included in the meta-regressions. The level of aerobic capacity is somehow affected by gender [22]. However, we did not find a correlation between gender distribution and the $ACLC_{diff}$. This makes our results applicable for future research and clinical use in single gender groups as well as mixed gender groups.

The $ACLC_{diff}$ does not seem to be affected by the risk of bias in the studies as low quality studies are reporting the same $ACLC_{diff}$ as high quality studies. This may be explained by the precise and accurate equipment used in CPET [61], and thereby the possibility of precise testing in different settings, which increases the clinical applicability.

The most accurate estimate of the ratio is the meta-analysis of the reported ratios, but only four studies [33,39,46,54] reported mean ± SD (%) values for the ratio between the tests. The meta-analysis revealed a linear relationship between the AC values and LC values with an $ACLC_{ratio}$ of 70%. This analysis should be seen as the main expression for the ratio between the values of the AC and the LC, where no important heterogeneity were found [62]. Three different methods were used to estimate the $ACLC_{ratio}$ due to the number of studies reporting values to incorporate in the meta-analysis for the ratio. The calculation and the linear regression of the $ACLC_{ratio}$ should only be used as a prediction, since they do not incorporate standard deviations. Despite different approaches to estimate the ratio, the results are very similar and the $ACLC_{ratio}$ of 70% is similar to the ones described in the literature [33,39,46,54]. To increase the power of this and investigate if the 70% is a valid estimate for the population mean $ACLC_{ratio}$, future research should report within comparison ratios between the AC and the LC, making them applicable for inclusion in meta-analysis.

This is the first systematic review and meta-analysis of literature comparing arm and leg exercise, and it is thus important to stress that our study has a number of limitations. First, some studies did not report $ACLC_{diff}$ standard deviation which meant we had to impute the value based on an assumed within participant correlation coefficient (r-value) between AC and LC $VO2_{max}$. This method is recommended by the Cochrane Handbook [21] but we acknowledge that it may influence the accuracy of our findings. The only way to avoid these limitations in a meta-analysis is for future research to report the correlation coefficients between the two tests. However, we undertook sensitivity analyses to assess the impact of this estimation on our findings. A small number of studies have reported a range of correlation coefficients between the AC test and the LC test (0.78, 0.94, 0.77, 0.32, 0.70) [12,17,31,37,54]. The pooled $ACLC_{diff}$ was found to be 12.52 ml/kg/min (95% CI: 10.2 to 14.6) based on the lowest of these r-values (0.32) and 12.6 ml/kg/min (95% CI: 10.6 to 14.7) with the highest reported r-value (0.94). In other words, this imputation method made little or no difference to the pooled results. Future studies need to report the standard deviation (or equivalent) of the mean difference between AC ad LC $VO_{2max}$ or the within person correlation coefficient.

Secondly, the quality of the included studies was variable. In this review, we sought to assess study risk of bias using the QAT, tool as it can be applied to cross-sectional studies [20]. However, to make this tool relevant to this review we had to adapt it by dropping some of the original QAT elements (items 6-10 and item 13)

Thirdly, this review was limited to non-athlete healthy adults and limits generalizability of our findings. Non-athlete healthy adults are expected to have a larger aerobic capacity when doing CPET using the legs compared to the arms due to everyday use and large lower limb muscle mass [29]. However, in athletic populations, particularly arm-trained populations, the $ACLC_{diff}$ is expected to be smaller than shown in this review [63]. To avoid systematic bias we excluded 18 comparisons in individuals performing more than 300 minutes per week of physical activity or involved in competitive exercise [19]. The groups contained ‘well trained subjects’, ‘triathletes’, ‘swimmers’, ‘cross-country skiers’ or ‘highly arm-trained’. However, we did not exclude studies in sedentary individuals. Two of the studies included extremely sedentary or sedentary subjects [39,47]. But having an $ACLC_{ratio}$ of 76% and 64% these studies are not likely to have had a systematic affect on our results. A sensitivity analysis was performed without the two studies and showed only minor impact on the result. The pooled $ACLC_{diff}$ was found to be 12.7 ml/kg/min (95% CI: 10.4 to 15.0). Future well conducted studies are needed that directly compare AC and LC in other populations, especially in disease populations with limitations by lower limb disability such as peripheral vascular disease or osteoarthritis.

## 5. Conclusion

This systematic review and meta-analysis showed that in studies on healthy non-athletic individuals although there was a linear association between the $VO_{2max}$ for AC and LC tests, the $VO_{2max}$ achieved by AC tests were on average 70% lower than compared to the LC. This magnitude of this difference appeared to be reduced in studies with older and less active populations.

## 6. Conflicts of interest

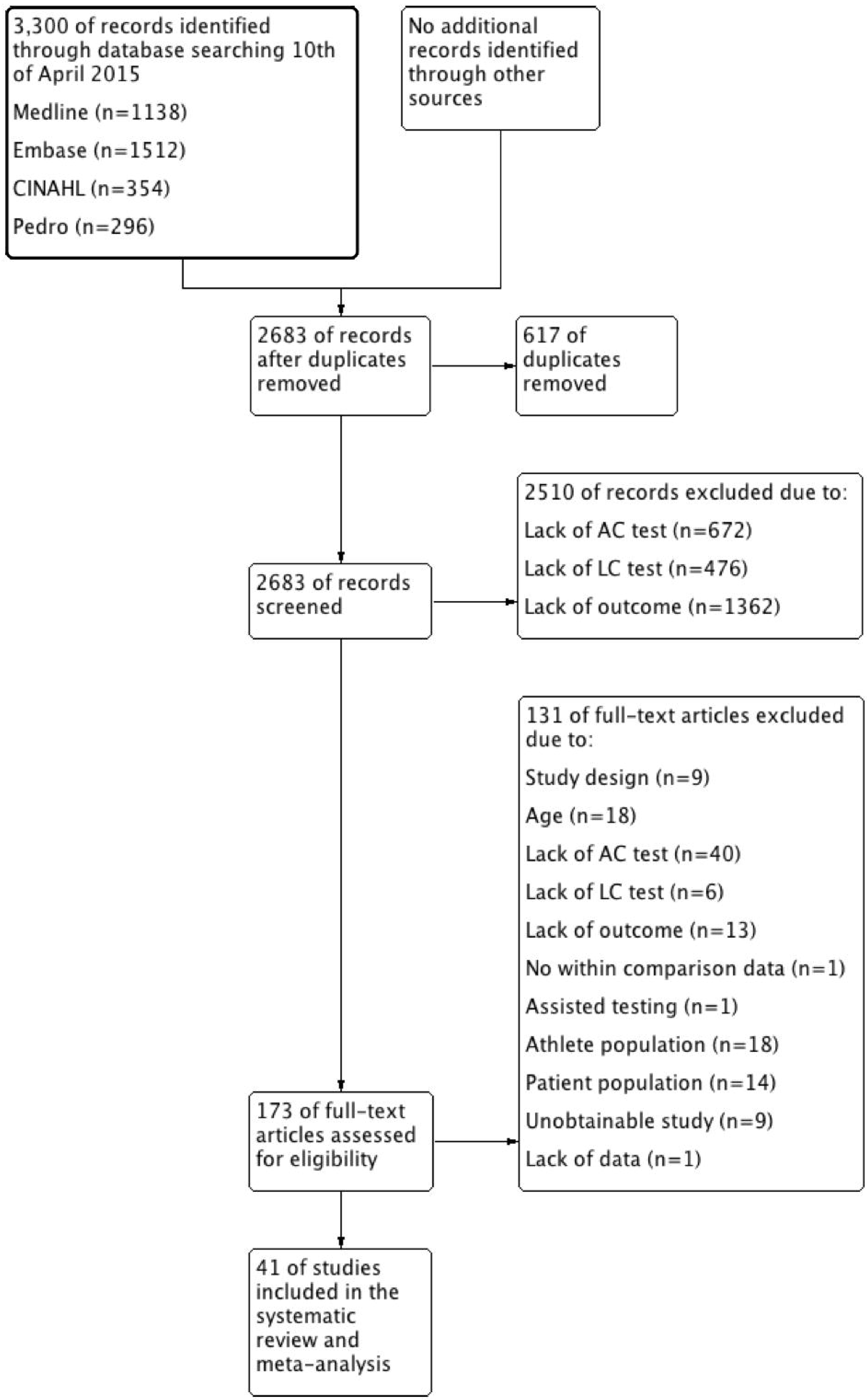
All authors declare that they have no conflict of interest.

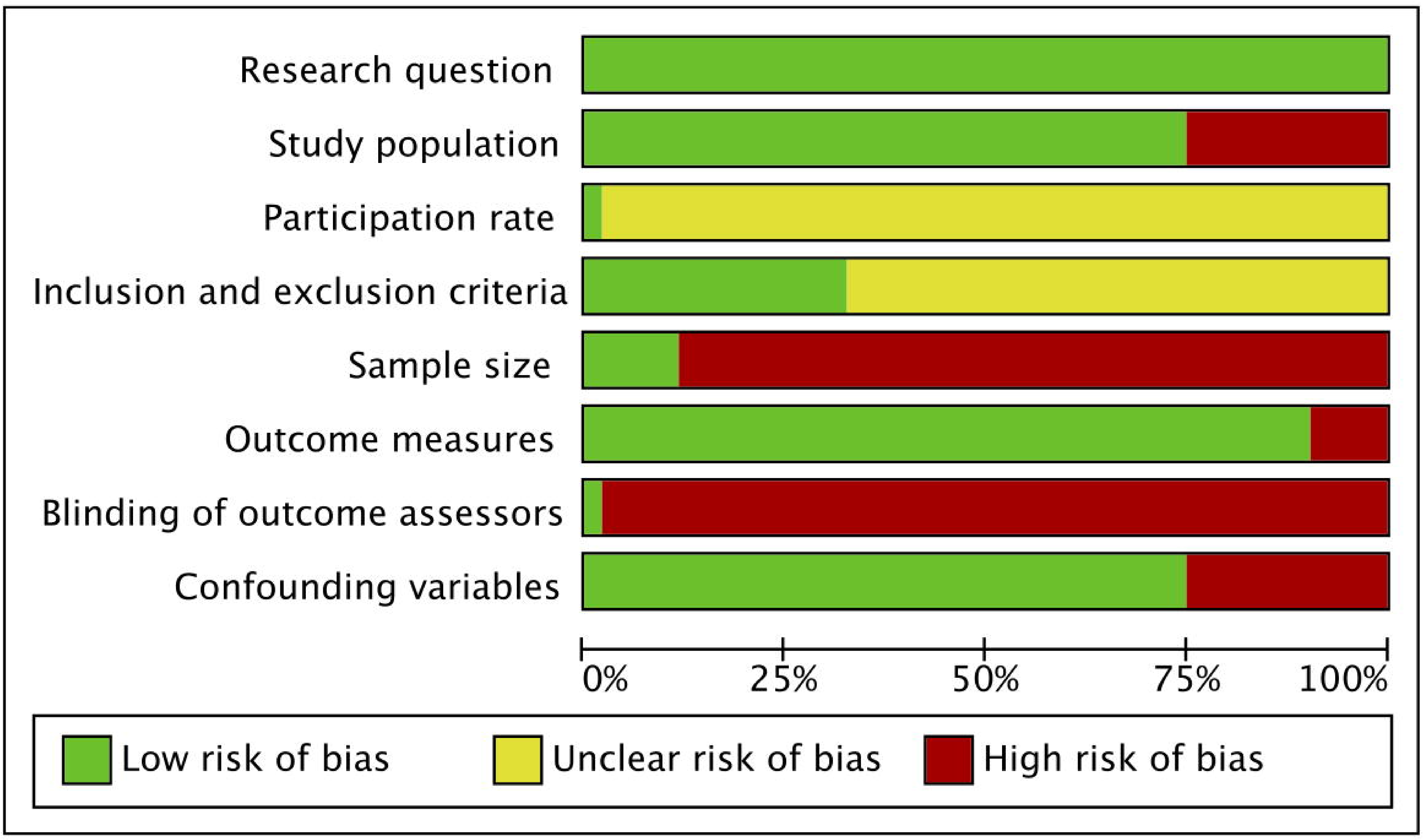
## References


1. Young EL, Karthikesalingam A, Huddart S, et al. A systematic review of the role of cardiopulmonary exercise testing in vascular surgery. *European journal of vascular and endovascular surgery : the official journal of the European Society for Vascular Surgery*. 2012;44(1):64-71.
2. Steins Bisschop CN, Velthuis MJ, Wittink H, et al. Cardiopulmonary exercise testing in cancer rehabilitation: a systematic review. *Sports medicine (Auckland, N.Z.)*. 2012;42(5):367-379.
3. Albouaini K, Egred M, Alahmar A. Cardiopulmonary exercise testing and its application. *Postgraduate Medical Journal*. 2007;83(985):675-682.
4. Herdy AH, Uhlendorf D. Reference values for cardiopulmonary exercise testing for sedentary and active men and women. *Arquivos brasileiros de cardiologia*. 2011;96(1):54-59.
5. Paap D, Takken T. Reference values for cardiopulmonary exercise testing in healthy adults: a systematic review. *Expert review of cardiovascular therapy*. 2014;12(12):1439-1453.
6. Stickland MK, Butcher SJ, Marciniuk DD, Bhutani M. Assessing Exercise Limitation Using Cardiopulmonary Exercise Testing. *Pulmonary Medicine*. 2012;2012:13.
7. ATS/ACCP Statement on cardiopulmonary exercise testing. *American journal of respiratory and critical care medicine*. 2003;167(2):211-277.
8. Mancini DM, Eisen H, Kussmaul W, Mull R, Edmunds LH, Jr., Wilson JR. Value of peak exercise oxygen consumption for optimal timing of cardiac transplantation in ambulatory patients with heart failure. *Circulation*. 1991;83(3):778-786.
9. Vanhees L, Fagard R, Thijs L, Amery A. Prognostic value of training-induced change in peak exercise capacity in patients with myocardial infarcts and patients with coronary bypass surgery. *The American journal of cardiology*. 1995;76(14):1014-1019.
10. West M, Jack S, Grocott MPW. Perioperative cardiopulmonary exercise testing in the elderly. *Best Practice & Research Clinical Anaesthesiology*. 2011;25(3):427-437.
11. Vanhees L, De Sutter J, Gelada SN, et al. Importance of characteristics and modalities of physical activity and exercise in defining the benefits to cardiovascular health within the general population: recommendations from the EACPR (Part I). *European journal of preventive cardiology*. 2012;19(4):670-686.
12. Loughney L, West M, Pintus S, et al. Comparison of oxygen uptake during arm or leg cardiopulmonary exercise testing in vascular surgery patients and control subjects. *British journal of anaesthesia*. 2014;112(1):57-65.
13. Walker RD, Nawaz S, Wilkinson CH, Saxton JM, Pockley AG, Wood RFM. Influence of upper- and lower-limb exercise training on cardiovascular function and walking distances in patients with intermittent claudication. *Journal of Vascular Surgery*. 2000;31(4):662-669.
14. Sutbeyaz ST, Sezer N, Koseoglu BF, Ibrahimoglu F, Tekin D. Influence of knee osteoarthritis on exercise capacity and quality of life in obese adults. *Obesity (Silver Spring, Md.)*. 2007;15(8):2071-2076.
15. Orr JL, Williamson P, Anderson W, Ross R, McCafferty S, Fettes P. Cardiopulmonary exercise testing: Arm crank vs cycle ergometry. *Anaesthesia*. 2013;68(5):497-501.
16. Secher NH, Volianitis S. Are the arms and legs in competition for cardiac output? *Med Sci Sports Exerc*. 2006;38(10):1797-1803.
17. Sawka MN, Foley ME, Pimental NA, Pandolf KB. Physiological factors affecting upper body aerobic exercise. *Ergonomics*. 1983;26(7):639-646.
18. Moher D, Liberati A, Tetzlaff J, Altman DG, The PG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med*. 2009;6(7):e1000097.
19. WHO. *Global recommendations on physical activity for health*. Switzerland: WHO;2010.
20. NHLBI. Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies. 2014; https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/cohort. Accessed 28.05.2015, 2015.
21. Higgins J, Deeks JJ, Altman DG. Chapter 16: Special topics in statistics. *Cochrane Handbook for Systematic Reviews of Interventions*. Vol 5.1.02011.
22. Åstrand I. Aerobic work capacity in men and women with special reference to age. *Acta physiologica Scandinavica. Supplementum*. 1960;49(169):1-92.

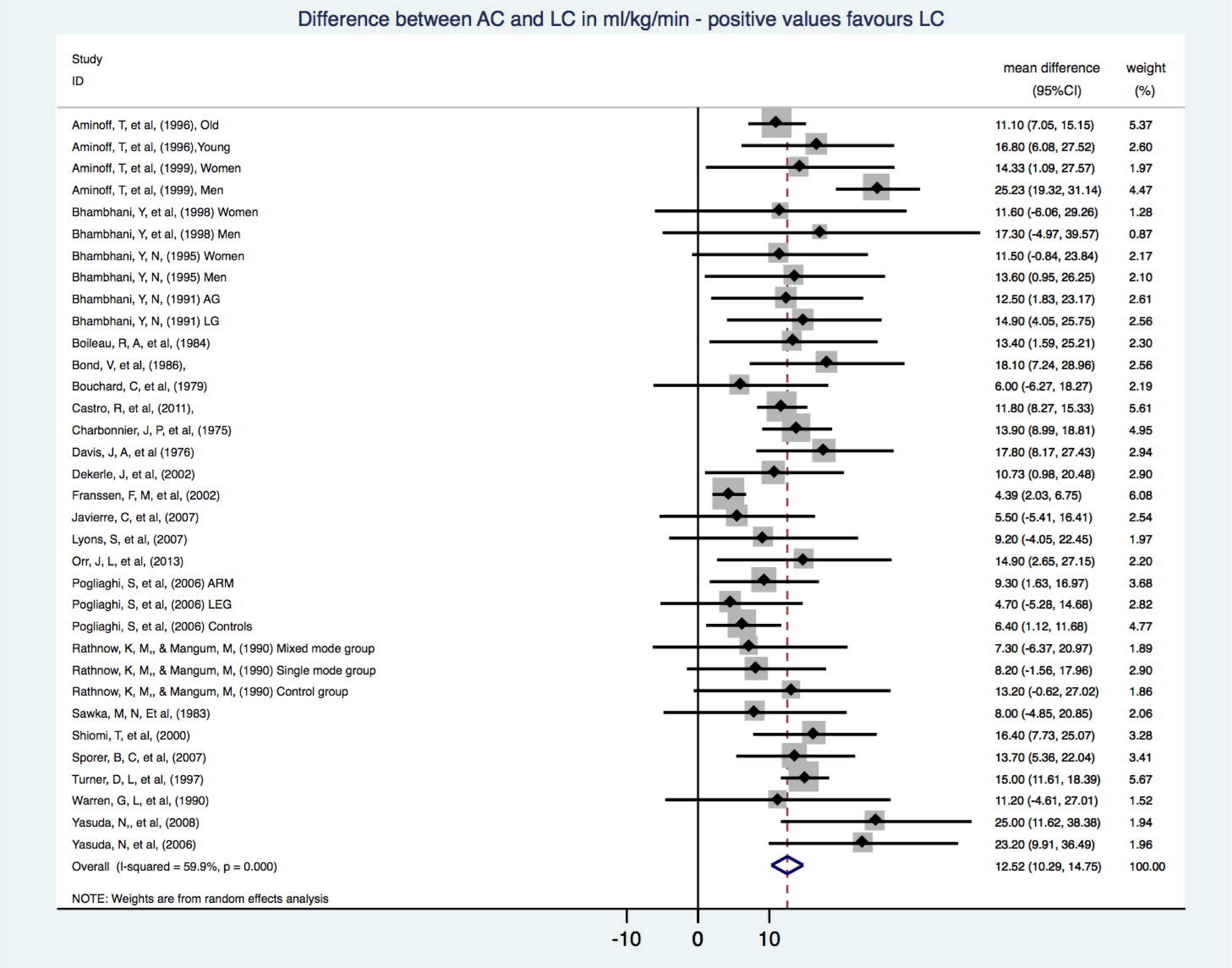
23. Aminoff T, Smolander J, Korhonen O, Louhevaara V. Physical work capacity in dynamic exercise with differing muscle masses in healthy young and older men. *Eur J Appl Physiol Occup Physiol.* 1996;73(1-2):180-185.
24. Aminoff T, Smolander J, Korhonen O, Louhevaara V. Physiological strain during kitchen work in relation to maximal and task-specific peak values. *Ergonomics.* 1999;42(4):584-592.
25. Barstow TJ, Casaburi R, Wasserman K. O2 uptake kinetics and the O2 deficit as related to exercise intensity and blood lactate. *Journal of Applied Physiology.* 1993;75(2):755-762.
26. Bhambhani Y, Maikala R, Buckley S. Muscle oxygenation during incremental arm and leg exercise in men and women. *Europ. J. Appl. Physiol.* 1998;78(5):422-431.
27. Bhambhani YN. Prediction of stroke volume during upper and lower body exercise in men and women. *Archives of physical medicine and rehabilitation.* 1995;76(8):713-718.
28. Bhambhani YN, Eriksson P, Gomes PS. Transfer effects of endurance training with the arms and legs. *Medicine and Science in Sports and Exercise.* 1991;23(9):1035-1041.
29. Boileau RA, McKeown BC, Riner WF. Cardiovascular and metabolic contributions to the maximal aerobic power of the arms and legs. *International Journal of Sports Cardiology.* 1984;1(2):67-75.
30. Bond V, Balkissoon B, Caprarola M, Tearney RJ. Aerobic capacity during two-arm and one-leg ergometric exercise. *International Rehabilitation Medicine.* 1986;8(2):79-81.
31. Bouchard C, Godbout P, Mondor JC, Leblanc C. Specificity of maximal aerobic power. *Europ. J. Appl. Physiol.* 1979;40(2):85-93.
32. Castro RRT, Pedrosa S, Nobrega ACL. Different ventilatory responses to progressive maximal exercise test performed with either the arms or legs. *Clinics.* 2011;66(7):1137-1142.
33. Charbonnier JP, Lacour JR, Riffat J, Flandrois R. Experimental study of the performance of competition swimmers. *Eur J Appl Physiol Occup Physiol.* 1975;34(3):157-167.
34. Davies CTM, Sargeant AJ. Indirect determination of maximal aerobic power output during work with one or two limbs. *Europ. J. Appl. Physiol.* 1974;32(3):207-215.
35. Davis JA, Vodak P, Wilmore JH. Anaerobic threshold and maximal aerobic power for three modes of exercise. *Journal of Applied Physiology.* 1976;41(4):544-550.
36. Dekerle J, Dupont L, Caby I, et al. Ventilatory thresholds in arm and leg exercises with spontaneously chosen crank and pedal rates. *Perceptual and motor skills.* 2002;95(3 Pt 2):1035-1046.
37. Franklin BA, Vander L, Wrisley D, Rubenfire M. Aerobic requirements of arm ergometry: Implications for exercise testing and training. *Physician and Sportsmedicine.* 1983;11(10):81-90.
38. Franssen FME, Wouters EFM, Baarends EM, Akkermans MA, Schols AMW. Arm mechanical efficiency and arm exercise capacity are relatively preserved in chronic obstructive pulmonary disease. *Medicine & Science in Sports & Exercise.* 2002;34(10):1570-1576.
39. Javierre C, Alegre J, Ventura JL, et al. Physiological responses to arm and leg exercise in women patients with chronic fatigue syndrome. *Journal of Chronic Fatigue Syndrome.* 2007;14(1):43-53.
40. Keteyian S, Marks CRC, Levine AB, et al. Cardiovascular responses of cardiac transplant patients to arm and leg exercise. *Europ. J. Appl. Physiol.* 1994;68(5):441-444.
41. Lewis S, Thompson P, Areskog NH, et al. Transfer effects of endurance training to exercise with untrained limbs. *Eur J Appl Physiol Occup Physiol.* 1980;44(1):25-34.
42. Louhevaara V, Sovijarvi A, Ilmarinen J, Teraslinna P. Differences in cardiorespiratory responses during and after arm crank and cycle exercise. *Acta Physiologica Scandinavica.* 1990;138(2):133-143.
43. Lyons S, Richardson M, Bishop P, Smith J, Heath H, Giesen J. Excess post-exercise oxygen consumption in untrained men following exercise of equal energy expenditure: Comparisons of upper and lower body exercise. *Diabetes, Obesity and Metabolism.* 2007;9(6):889-894.
44. McConnell TR, Swett DD, Jeresaty RM. The hemodynamic and physiologic differences between exercise modalities. *Journal of Sports Medicine and Physical Fitness.* 1984;24(3):238-245.
45. Nag PK. Circulo-respiratory responses to different muscular exercises. *Europ. J. Appl. Physiol.* 1984;52(4):393-399.
46. Pogliaghi S, Terziotti P, Cevese A, Balestreri F, Schena F. Adaptations to endurance training in the healthy elderly: Arm cranking versus leg cycling. *European Journal of Applied Physiology.* 2006;97(6):723-731.
47. Protas EJ, Stanley RK, Jankovic J, MacNeill B. Cardiovascular and metabolic responses to upper- and lower-extremity exercise in men with idiopathic Parkinson's disease. *Physical Therapy.* 1996;76(1):34-40.
48. Ramonatxo M, Prioux J, Prefaut C. Differences in mouth occlusion pressure and breathing pattern between arm and leg incremental exercise. *Acta Physiologica Scandinavica.* 1996;158(4):333-341.
49. Rathnow KM, Mangum M. A comparison of single-versus multi-modal exercise programs: Effects on aerobic power. *Journal of Sports Medicine and Physical Fitness.* 1990;30(4):382-388.

50. Reybrouck T, Heigenhauser GF, Faulkner JA. Limitations to maximum oxygen uptake in arm, leg, and combined arm leg ergometry. *Journal of Applied Physiology.* 1975;38(5):774-779.
51. Rosler K, Hoppeler H, Conley KE, Claassen H, Gehr P, Howald H. Transfer effects in endurance exercise. Adaptations in trained and untrained muscles. *Eur J Appl Physiol Occup Physiol.* 1985;54(4):355-362.
52. Sargeant AJ, Davies CTM. Perceived exertion during rhythmic exercise involving different muscle masses. *Journal of Human Ergology.* 1973;2(1):3-11.
53. Sharp MA, Harman E, Vogel JA, Knapik JJ, Legg SJ. Maximal aerobic capacity for repetitive lifting: Comparison with three standard exercise testing modes. *Europ. J. Appl. Physiol.* 1988;57(6):753-760.
54. Shiomi T, Maruyama H, Saito A, Umemura M. Physiological responses and mechanical efficiency during different types of ergometric exercise. *Journal of Physical Therapy Science.* 2000;12(1):67-73.
55. Sporer BC, Foster GE, Sheel AW, McKenzie DC. Entrainment of breathing in cyclists and non-cyclists during arm and leg exercise. *Respiratory Physiology and Neurobiology.* 2007;155(1):64-70.
56. Swensen TC, Howley ET. Effect of one- and two-leg training on arm and two-leg maximum aerobic power. *Eur J Appl Physiol Occup Physiol.* 1993;66(3):285-288.
57. Turner DL, Hoppeler H, Claassen H, et al. Effects of endurance training on oxidative capacity and structural composition of human arm and leg muscles. *Acta Physiol Scand.* 1997;161(4):459-464.
58. Warren GL, Cureton KJ, Dengel DR, Graham RE, Ray CA. Is the gender difference in peak V(O2) greater for arm than leg exercise? *Europ. J. Appl. Physiol.* 1990;60(2):149-154.
59. Yasuda N, Gaskill SE, Ruby BC. No gender-specific differences in mechanical efficiency during arm or leg exercise relative to ventilatory threshold. *Scandinavian Journal of Medicine and Science in Sports.* 2008;18(2):205-212.
60. Yasuda N, Ruby BC, Gaskill SE. Substrate oxidation during incremental arm and leg exercise in men and women matched for ventilatory threshold. *Journal of Sports Sciences.* 2006;24(12):1281-1289.
61. Bhagwat M, Paramesh K. Cardio-pulmonary exercise testing: An objective approach to pre-operative assessment to define level of perioperative care. *Indian Journal of Anaesthesia.* 2010;54(4):286-291.
62. Higgins J, Deeks JJ, Altman DG. 9.5.2 Identifying and measuring heterogeneity. *Cochrane Handbook for Systematic Reviews of Interventions.* Vol 5.1.02011.
63. Secher NH, Ruberg-Larsen N, Binkhorst RA, Bonde-Petersen F. Maximal oxygen uptake during arm cranking and combined arm plus leg exercise. *J Appl Physiol.* 1974;36(5):515-518.

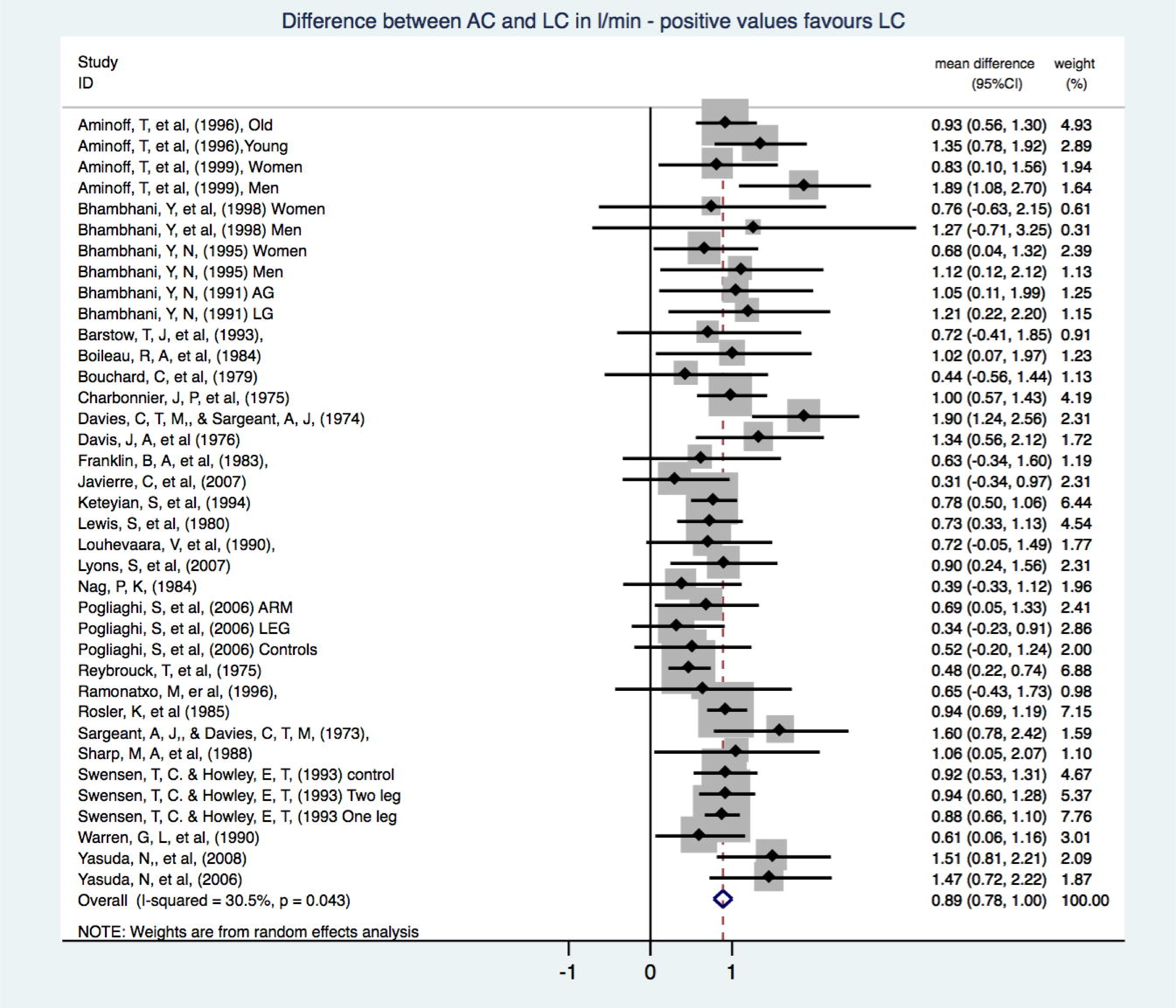

3,300 of records identified through database searching 10th of April 2015
Medline (n=1138)
Embase (n=1512)
CINAHL (n=354)
Pedro (n=296)
No additional records identified through other sources
2683 of records after duplicates removed
617 of duplicates removed
2510 of records excluded due to:
Lack of AC test (n=672)
Lack of LC test (n=476)
Lack of outcome (n=1362)
2683 of records screened
131 of full-text articles excluded due to:
Study design (n=9)
Age (n=18)
Lack of AC test (n=40)
Lack of LC test (n=6)
Lack of outcome (n=13)
No within comparison data (n=1)
Assisted testing (n=1)
Athlete population (n=18)
Patient population (n=14)
Unobtainable study (n=9)
Lack of data (n=1)
173 of full-text articles assessed for eligibility
41 of studies included in the systematic review and meta-analysis

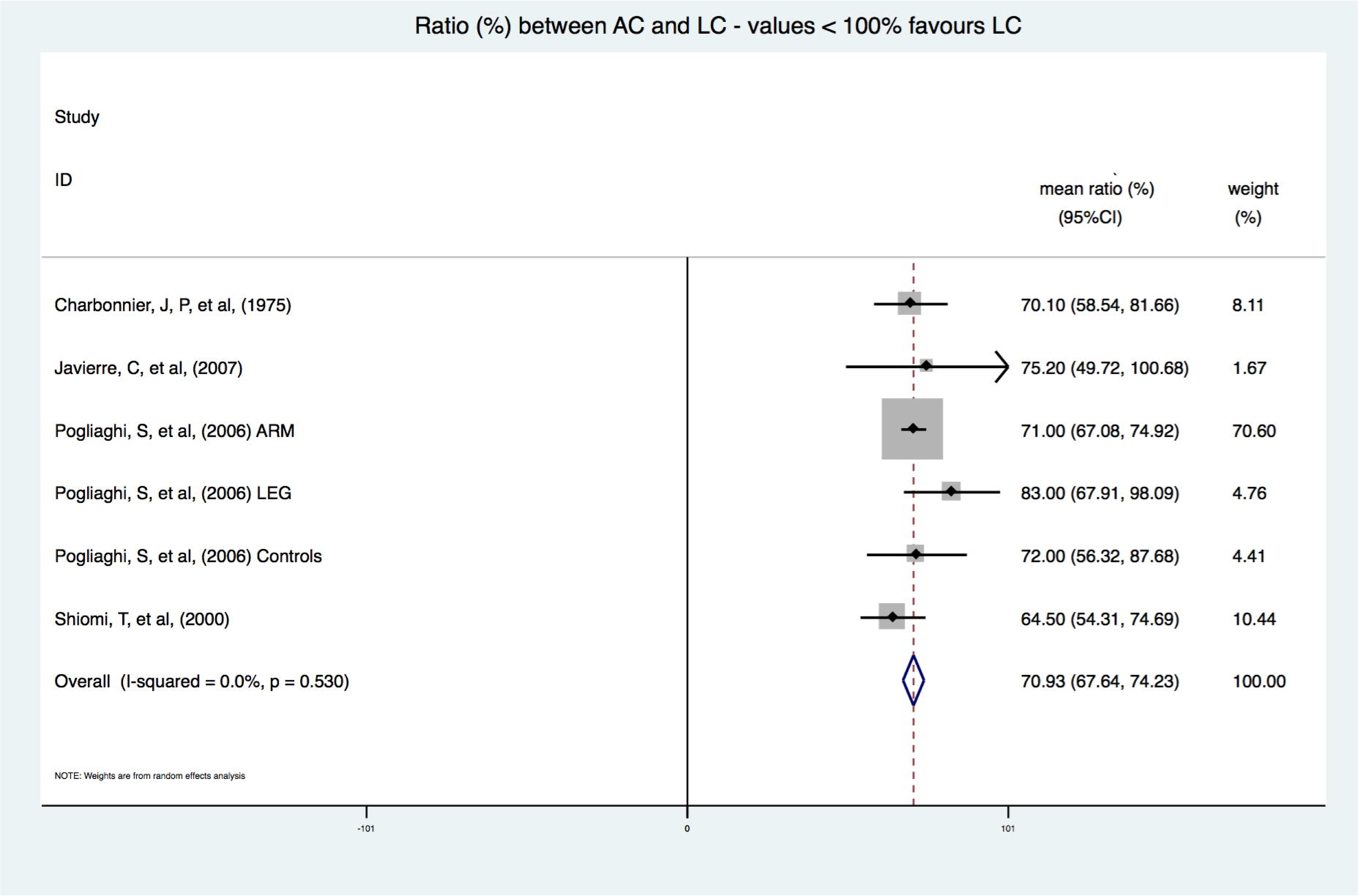


Research question
Study population
Participation rate
Inclusion and exclusion criteria
Sample size
Outcome measures
Blinding of outcome assessors
Confounding variables
0%
25%
50%
75%
100%
Low risk of bias
Unclear risk of bias
High risk of bias

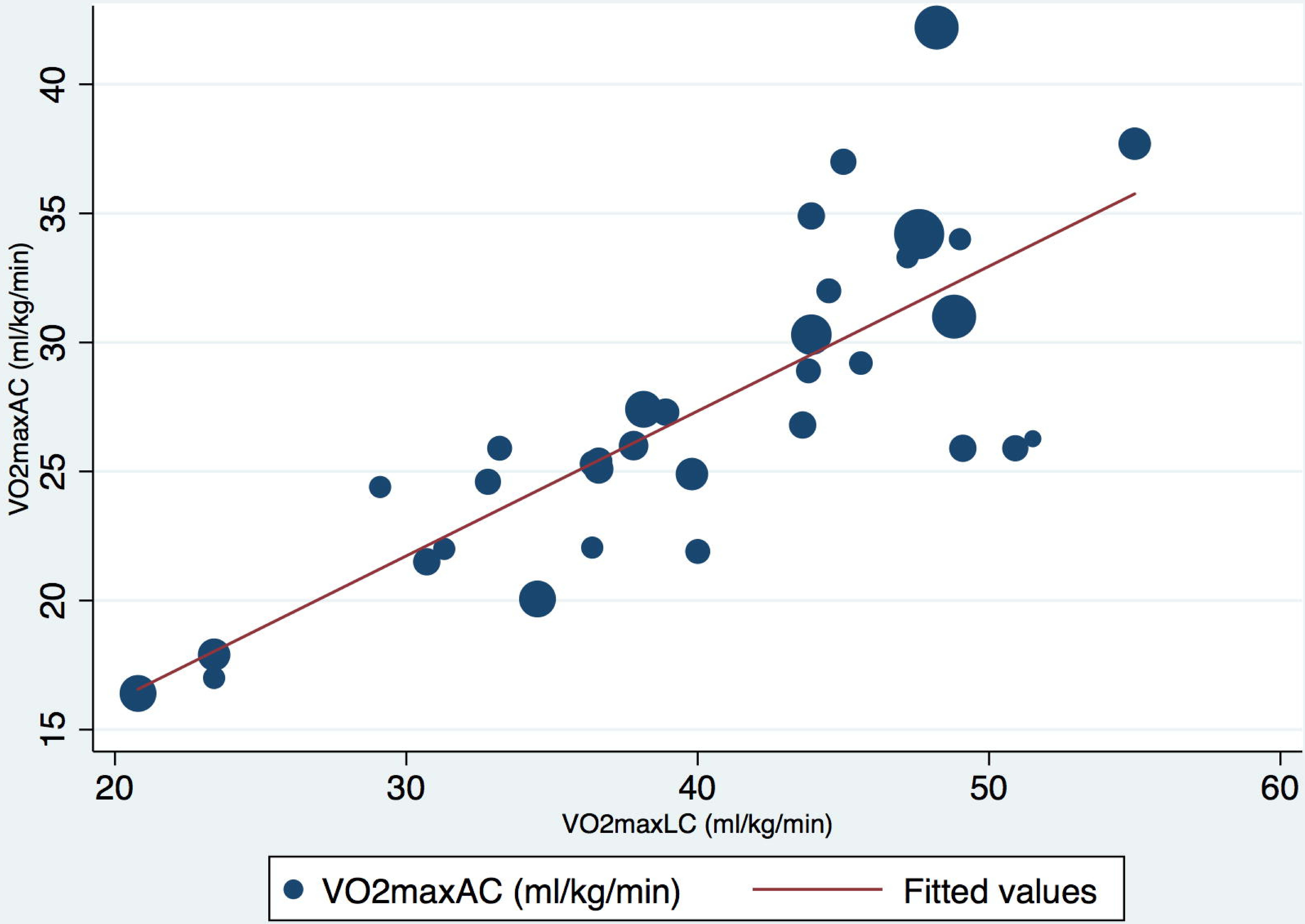


Difference between AC and LC in ml/kg/min - positive values favours LC
Study
ID
mean difference
(95%CI)
weight
(%)
Aminoff, T, et al, (1996), Old
11.10 (7.05, 15.15)
5.37
Aminoff, T, et al, (1996),Young
16.80 (6.08, 27.52)
2.60
Aminoff, T, et al, (1999), Women
14.33 (1.09, 27.57)
1.97
Aminoff, T, et al, (1999), Men
25.23 (19.32, 31.14)
4.47
Bhambhani, Y, et al, (1998) Women
11.60 (-6.06, 29.26)
1.28
Bhambhani, Y, et al, (1998) Men
17.30 (-4.97, 39.57)
0.87
Bhambhani, Y, N, (1995) Women
11.50 (-0.84, 23.84)
2.17
Bhambhani, Y, N, (1995) Men
13.60 (0.95, 26.25)
2.10
Bhambhani, Y, N, (1991) AG
12.50 (1.83, 23.17)
2.61
Bhambhani, Y, N, (1991) LG
14.90 (4.05, 25.75)
2.56
Boileau, R, A, et al, (1984)
13.40 (1.59, 25.21)
2.30
Bond, V, et al, (1986),
18.10 (7.24, 28.96)
2.56
Bouchard, C, et al, (1979)
6.00 (-6.27, 18.27)
2.19
Castro, R, et al, (2011),
11.80 (8.27, 15.33)
5.61
Charbonnier, J, P, et al, (1975)
13.90 (8.99, 18.81)
4.95
Davis, J, A, et al (1976)
17.80 (8.17, 27.43)
2.94
Dekerle, J, et al, (2002)
10.73 (0.98, 20.48)
2.90
Franssen, F, M, et al, (2002)
4.39 (2.03, 6.75)
6.08
Javierre, C, et al, (2007)
5.50 (-5.41, 16.41)
2.54
Lyons, S, et al, (2007)
9.20 (-4.05, 22.45)
1.97
Orr, J, L, et al, (2013)
14.90 (2.65, 27.15)
2.20
Pogliaghi, S, et al, (2006) ARM
9.30 (1.63, 16.97)
3.68
Pogliaghi, S, et al, (2006) LEG
4.70 (-5.28, 14.68)
2.82
Pogliaghi, S, et al, (2006) Controls
6.40 (1.12, 11.68)
4.77
Rathnow, K, M,, & Mangum, M, (1990) Mixed mode group
7.30 (-6.37, 20.97)
1.89
Rathnow, K, M,, & Mangum, M, (1990) Single mode group
8.20 (-1.56, 17.96)
2.90
Rathnow, K, M,, & Mangum, M, (1990) Control group
13.20 (-0.62, 27.02)
1.86
Sawka, M, N, Et al, (1983)
8.00 (-4.85, 20.85)
2.06
Shiomi, T, et al, (2000)
16.40 (7.73, 25.07)
3.28
Sporer, B, C, et al, (2007)
13.70 (5.36, 22.04)
3.41
Turner, D, L, et al, (1997)
15.00 (11.61, 18.39)
5.67
Warren, G, L, et al, (1990)
11.20 (-4.61, 27.01)
1.52
Yasuda, N,, et al, (2008)
25.00 (11.62, 38.38)
1.94
Yasuda, N, et al, (2006)
23.20 (9.91, 36.49)
1.96
Overall (I-squared = 59.9%, p = 0.000)
12.52 (10.29, 14.75)
100.00
NOTE: Weights are from random effects analysis
-10
0
10


Difference between AC and LC in l/min - positive values favours LC
Study
ID
mean difference
(95%CI)
weight
(%)
Aminoff, T, et al, (1996), Old 0.93 (0.56, 1.30) 4.93
Aminoff, T, et al, (1996),Young 1.35 (0.78, 1.92) 2.89
Aminoff, T, et al, (1999), Women 0.83 (0.10, 1.56) 1.94
Aminoff, T, et al, (1999), Men 1.89 (1.08, 2.70) 1.64
Bhambhani, Y, et al, (1998) Women 0.76 (-0.63, 2.15) 0.61
Bhambhani, Y, et al, (1998) Men 1.27 (-0.71, 3.25) 0.31
Bhambhani, Y, N, (1995) Women 0.68 (0.04, 1.32) 2.39
Bhambhani, Y, N, (1995) Men 1.12 (0.12, 2.12) 1.13
Bhambhani, Y, N, (1991) AG 1.05 (0.11, 1.99) 1.25
Bhambhani, Y, N, (1991) LG 1.21 (0.22, 2.20) 1.15
Barstow, T, J, et al, (1993), 0.72 (-0.41, 1.85) 0.91
Boileau, R, A, et al, (1984) 1.02 (0.07, 1.97) 1.23
Bouchard, C, et al, (1979) 0.44 (-0.56, 1.44) 1.13
Charbonnier, J, P, et al, (1975) 1.00 (0.57, 1.43) 4.19
Davies, C, T, M,, & Sargeant, A, J, (1974) 1.90 (1.24, 2.56) 2.31
Davis, J, A, et al (1976) 1.34 (0.56, 2.12) 1.72
Franklin, B, A, et al, (1983), 0.63 (-0.34, 1.60) 1.19
Javierre, C, et al, (2007) 0.31 (-0.34, 0.97) 2.31
Keteyian, S, et al, (1994) 0.78 (0.50, 1.06) 6.44
Lewis, S, et al, (1980) 0.73 (0.33, 1.13) 4.54
Louhevaara, V, et al, (1990), 0.72 (-0.05, 1.49) 1.77
Lyons, S, et al, (2007) 0.90 (0.24, 1.56) 2.31
Nag, P, K, (1984) 0.39 (-0.33, 1.12) 1.96
Pogliaghi, S, et al, (2006) ARM 0.69 (0.05, 1.33) 2.41
Pogliaghi, S, et al, (2006) LEG 0.34 (-0.23, 0.91) 2.86
Pogliaghi, S, et al, (2006) Controls 0.52 (-0.20, 1.24) 2.00
Reybrouck, T, et al, (1975) 0.48 (0.22, 0.74) 6.88
Ramonatxo, M, er al, (1996), 0.65 (-0.43, 1.73) 0.98
Rosler, K, et al (1985) 0.94 (0.69, 1.19) 7.15
Sargeant, A, J,, & Davies, C, T, M, (1973), 1.60 (0.78, 2.42) 1.59
Sharp, M, A, et al, (1988) 1.06 (0.05, 2.07) 1.10
Swensen, T, C. & Howley, E, T, (1993) control 0.92 (0.53, 1.31) 4.67
Swensen, T, C. & Howley, E, T, (1993) Two leg 0.94 (0.60, 1.28) 5.37
Swensen, T, C. & Howley, E, T, (1993 One leg 0.88 (0.66, 1.10) 7.76
Warren, G, L, et al, (1990) 0.61 (0.06, 1.16) 3.01
Yasuda, N,, et al, (2008) 1.51 (0.81, 2.21) 2.09
Yasuda, N, et al, (2006) 1.47 (0.72, 2.22) 1.87
Overall (I-squared = 30.5%, p = 0.043) 0.89 (0.78, 1.00) 100.00
NOTE: Weights are from random effects analysis
-1
0
1



Ratio (%) between AC and LC - values < 100% favours LC
Study
ID
mean ratio (%)
(95%CI)
weight
(%)
Charbonnier, J, P, et al, (1975)
70.10 (58.54, 81.66)
8.11
Javierre, C, et al, (2007)
75.20 (49.72, 100.68)
1.67
Pogliaghi, S, et al, (2006) ARM
71.00 (67.08, 74.92)
70.60
Pogliaghi, S, et al, (2006) LEG
83.00 (67.91, 98.09)
4.76
Pogliaghi, S, et al, (2006) Controls
72.00 (56.32, 87.68)
4.41
Shiomi, T, et al, (2000)
64.50 (54.31, 74.69)
10.44
Overall (I-squared = 0.0%, p = 0.530)
70.93 (67.64, 74.23)
100.00
NOTE: Weights are from random effects analysis
-101
0
101



40
35
30
25
20
15
VO2maxAC (ml/kg/min)
20
30
40
50
60
VO2maxLC (ml/kg/min)
VO2maxAC (ml/kg/min)
Fitted values

**Table 1 -** Study Characteristics of the 53 groups from the 41 included studies

| **Continent of publication** | **(%)** |
|---|---|
| North America | 56.6 % |
| Europe | 35.8 % |
| South America | 3.8 % |
| Asia | 3.8 % |
| | |
| **Study Design** | |
| RCT | 17.0 % |
| Non-RCT | 3.8 % |
| Cross-sectional | 79.2 % |
| | |
| **Study risk of bias** | **Median (IQR)** |
| SumQAT | 4 points (3 to 5) |
| | |
| **Participant characteristics** | |
| | |
| **Gender** | **(%)** |
| Male only | 66 % |
| Female only | 15.1 % |
| Mixed | 15.1 % |
| Not reported | 3.8 % |
| | |
| | **Median (IQR)** |
| Mean age years | 28.4 years (25 to 32.3) |
| Mean BMI, kg/m$^2$ | 23.65 kg/m$^2$ (22.7 to 25) |
| | |
| **Aerobic capacity** | **(%)** |
| Low | 3.8 % |
| Average | 28.7 % |
| Good | 5.6 % |
| High | 3.8 % |
| Did not report | 58.1 % |
| | |
| **Test characteristics** | |
| | |
| Order on AC/LC test | **(%)** |
| AC first | 3.8 % |
| LC first | 18.9 % |
| Random order | 45.3 % |
| Not reported | 32 % |
| | |
| | **Median (IQR)** |
| Time between tests (hours) | 72 (24 to 168) |
| | |
| AC start level (watts) | 25 (15 to 40) |
| LC start levels (watts) | 50 (30 to 50) |
| | |
| AC increase/min (watt) | 10.7 (5 to 17) |
| LC increase/min (watt) | 30 (20.7 to 30) |

IQR: Interquartile range, SumQAT: sum of quality assessment tool score, AC: Arm cycle, LC: Leg cycle

***Table 2*** **–** Meta-regression analyses performed on each variable (univariable) and adjusted for all variables (multivariable)

| **Univariable meta-regression on $ACLC_{diff}$** | **Groups included in analysis** | **Mean coefficient (95% CI)** | **p-value** |
|---|---|---|---|
| Aerobic capacity | 27 | 4.1 (95% CI: 1.5 to 6.6) | p=0.003 |
| Gender distribution (% male) | 33 | -1.25 (95% CI: -7.4 to 4.9) | p=0.684 |
| Mean age | 29 | -2.1 (95% CI: -0.3 to -0.1) | p<0.001 |
| Mean difference in peak RER values | 24 | -12.1 (95% CI: -68.8 to 44.6) | p=0.663 |
| Risk of bias (SumQAT score) | 34 | -0.19 (95% CI: -2.6 to 2.2) | p=0.875 |
| **Multivariable meta-regression on $ACLC_{diff}$** | | | |
| Aerobic capacity | 16 | 4.0 (95% CI: 0.81 to 7.2) | p=0.019 |
| Gender distribution (% male) | 16 | 4.5 (95% CI: -4.1 to 13.2) | p=0.268 |
| Mean age (years) | 16 | -0.25 (95% CI: -0.4 to -0.06) | p=0.014 |
| Mean difference in peak RER values | 16 | 7.9 (95% CI: -59.0 to 74.8) | p=0.797 |
| Risk of bias (SumQAT score) | 16 | 0.9 (95% CI: -3.8 to 5.6) | p=0.682 |

$ACLC_{diff}$: difference between obtained AC $VO_{2max}$ and obtained LC $VO_{2max}$, 95% CI: 95% confidence intervals, RER: respiratory exchange ratio, SumQAT: sum of quality assessment tool score